DOCUMENT RESUME

ED 400 102 PS 024 647

AUTHOR	Friedman, Mark
TITLE	A Strategy Map for Results-Based Budgeting: Moving from Theory to Practice.
INSTITUTION	Finance Project, Washington, DC.
PUB DATE	Sep 96
NOTE	61p.
PUB TYPE	Guides - Non-Classroom Use (055)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	*Budgeting; *Budgets; *Educational Finance; Models; *Money Management; Organizational Effectiveness; Organizational Objectives; *Program Budgeting; Program Improvement; Statewide Planning; *Theory Practice Relationship; Total Quality Management
IDENTIFIERS	Continuous Quality Improvement

ABSTRACT

Public financing for education and an array of other children's services has become a topic of significant interest and political concern. Elected officials and policymakers have sought to respond to public concern and dissatisfaction through a more explicit focus on results. To assist in this work of achieving and measuring important outcomes for children and their families, the Finance Project created a Working Group on Results-based Planning, Budgeting, Management, and Accountability Systems. This paper is a result of the Working Group's efforts to design a strategic plan for those engaged in results-based budgeting. The first part of the paper provides definitions for some of the essential concepts needed for discussing results-based decision-making and budgeting. The themes covered include definitions for outcomes, indicators, and performance measures as well as a framework for choosing them. The second part suggests a strategy map, describing each element of the results-based budgeting strategy and proposing ideas about how to put these elements in place. In addition, the strategic map suggests how to build political and community support, how to reallocate resources and tie them to results, how to integrate results-based budgeting into the existing budget process, and how to avoid common pitfalls. The third part discusses lessons from state, local, and private initiatives to define, measure, and achieve results, and summarizes what has been learned so far. The concluding part stresses the importance for changes in the public budgeting system and highlights the role of results-based budgeting. (AA)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 400 102

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

A STRATEGY MAP FOR RESULTS-BASED BUDGETING

■

Moving from Theory to Practice

PS 024647

PERMISSION TO REPRODUCE AND
DISSEMINATE THIS MATERIAL
HAS BEEN GRANTED BY
C. D. Hayes
TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)

THE FINANCE PROJECT

BEST COPY AVAILABLE

A STRATEGY MAP FOR RESULTS-BASED BUDGETING

■

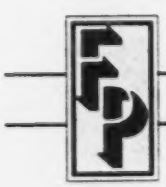
Moving from Theory to Practice

September 1996

By Mark Friedman

Prepared for

THE FINANCE PROJECT
1341 G Street, NW
Washington, DC 20005
202-628-4200
Fax: 202-628-4205



THE FINANCE PROJECT

ABOUT THE AUTHOR

Mark Friedman is an independent consultant, formerly with the Center for the Study of Social Policy. He is a member of The Finance Project's Working Group on Results-Based Planning, Budgeting, Management and Accountability Systems.

TABLE OF CONTENTS

Preface	i
Introduction	1
Starting Points	3
A Strategy Map for Results-Based Budgeting	10
Lessons and Issues	41
Conclusion	52

PREFACE

Public financing for education and an array of other children's services has become a topic of significant interest and political concern. Growing skepticism among a critical mass of American voters and taxpayers has fueled doubts about the ability of government to solve problems and provide basic supports and services that enhance the quality of life in their communities. Many believe government is too big, that it's too expensive, and that it doesn't work very well.

Despite steadily increasing public expenditures for health, education, welfare, human services, and public safety over the past two decases, seemingly intractable problems persist. Nearly a quarter of U.S. children are poor and live in families and communities that are unable to meet their basic needs. Schools have become increasingly expensive, but student achievement has not matched the rising costs and dropout rates remain unacceptably high. Health care costs continue to go up, yet many Americans can't get the services they need, and with each passing year their health care dollars buy less. Criminal justice demands a dramatically increasing share of public dollars -- for police officers and judges and jails -- but neighborhood streets don't seem any safer.

Voters have spoken clearly. They want more for their money. They have called for more and better services, but they also have demanded balanced budgets and cuts in income and property taxes. After more than a decade of chronic deficits, they want government at all levels to operate more effectively and efficiently. They don't want to dismantle government, but rather they want government to meet vital public needs and make a more visible difference in their lives.

Elected officials and other policymakers have sought to respond to public concern and dissatisfaction through a more explicit focus on results. Reformers have sought to redefine the missions of public programs and agencies, to modify how services are delivered, to measure how well government programs and agencies are performing, and to feed information about performance back into planning, budgeting, management, and accountability systems. While the

federal government's National Performance Review and its initiatives to "reinvent government" may be the most prominent example of the new focus on results, there are countless others at the state and local levels that span the divisions of ideology, political party, and the executive and legislative branches of government.

This focus on results is particularly important for programs and policies serving children and their families. The future well-being of the nation is obviously tied to the healthy development of children. Yet policymakers and citizens alike may be inclined to reduce their commitment to supports and services for children without stronger evidence that these efforts yield results that society cares about, such as quality education and health care, and safe homes and neighborhoods.

Unfortunately, many of the efforts to implement a results framework -- for children and their families as well as for public programs in general -- have been marred by confusion about terms and basic definitions, a lack of political understanding and support, the difficulty of defining results and appropriate measures, and the challenges of overhauling existing budget and management systems to emphasize results. Policymakers trying to implement results-based systems have enthusiastically set out in many different directions, but often without a particular destination in mind or a map to help them get there.

To assist with this important work of achieving and measuring important outcomes for children and their families, The Finance Project created a Working Group on Results-Based Planning, Budgeting, Management, and Accountability Systems. Established by a consortium of national foundations, The Finance Project has been conducting an ambitious agenda of policy research and development activities to improve the effectiveness, efficiency, and equity of public financing for education and other children's services.

This paper, *A Strategy Map for Results-Based Budgeting: Moving from Theory to Practice*, reflects the efforts of the Working Group on Results-based Planning, Budgeting, Management, and Accountability Systems to design a road map for those embarking on the path of results-based budgeting. Prepared by Mark Friedman, a member of the Working Group, the paper defines results, indicators, and performance measures and offers a framework for choosing them. The strategy map also discusses lessons from state, local, and private initiatives to define, measure, and achieve results. Furthermore, it suggests how to build political and community support, how to reallocate resources and tie them to results, how to integrate results-based budgeting into the existing budget process, and how to avoid common pitfalls. Finally, it will serve as a framework for a series of papers and tool kits by The Finance Project, including a guidebook on results and indicators, a guidebook on performance measures, a tool kit on children's budgets, and a paper

on cost-of bad-results analysis. These forthcoming products will expand on the concepts introduced in the strategy map. The strategy map was reviewed by Charles Bruner, Sid Gardner, and Cornelius D. Hogan. Their helpful and constructive comments are reflected in the final version.

Cheryl D. Hayes
Executive Director

INTRODUCTION

The concept of results-based budgeting is simple and literally business-like: Start with the results we want for children, families, and communities and work backward to the means to achieve those results. But how do we translate this simple concept into practice in the complex environment of public decision-making and budgeting?

A growing number of states, counties, cities, and communities are engaged in the work of identifying the results they want for children and families. In some cases, these efforts focus on matters of family and child well-being; in other cases, they concentrate on a more broadly based articulation of the desired quality of life for all citizens. But the challenge in each case is the same: to get from talking about results actually to doing something about them. This paper (and its companions) attempt to answer this central "talk-to-action" question. If results are things that matter to the long-term well-being of our society, how do we connect them to the work of actually deciding on our course of action and use of resources?

This paper is the sequel to a paper completed in draft in July 1995 by the Center for the Study of Social Policy. "From Outcomes to Budgets"[1] presents a conceptual framework for results-based budgeting and some beginning ideas about how to put such an approach in place.[2] The paper that follows is intended to take the next step, namely, to present in more detail how we believe such a system of results-based budgeting and decision-making could be implemented, and to illustrate implementation steps with the experiences of states and localities that are beginning to experiment with this approach.

It is important to note, at the beginning, that this is difficult and challenging work, and that there are no fully implemented results-based systems in place with all of the elements discussed in this paper. There is a lot of good work to show, but there are no complete systems ready to transfer or replicate.[3] The sections that follow describe some of what has been learned about

---

[1]The terms "outcomes" and "results" can be used interchangeably. In this paper, the term "results" is given preference over the term "outcomes."

[2] A related paper, "Trading Outcome Accountability for Fund Flexibility," available from the Center for the Study of Social Policy, Washington, D.C., explores how outcomes can be used in negotiating new state/local (fiscal) agreements, thereby providing more flexible use of categorical service dollars to achieve improved outcomes for families and children.

[3]As noted below, the distinction between results-based and performance-based budgeting is crucial. See the definitions in section II and the discussion in section IV. If the reader is interested only in performance-based systems, there is a large body of work in this area that can be studied and possibly replicated.

these systems over the past several years, and present a road map (or what we will call a strategy map) to full implementation of a results-based decision-making and budgeting system.

A word about confidentiality is in order. It is important in this work to give credit to the people who are actually doing the work, and we will try to identify by name those states and localities in the lead. But we will also be addressing some of the problems that come up in developing results-based systems. It is difficult to do this if we have to name the specific places involved, yet it is important to convey that the lessons are real and not hypothetical. We will use a convention, therefore, that allows us to refer to specific state and local issues without naming the jurisdictions.

II. STARTING POINTS:

This section provides, in abbreviated form, some of the essential starting points for a discussion of results-based decision-making and budgeting. If you have already read "From Outcomes to Budgets," then this section may be skipped.

A. Common Language

The choice is between discipline and the Tower of Babel. There is an astounding lack of discipline in the use of language in the current work being done on child and family well-being. It is quite common to find people working on these problems who use the same terms in different, sometimes contradictory ways, and wonder why they are not making any progress. Processes without a common language tend to be frustrating and ineffective. The work becomes mired. Attention and political energies focus elsewhere. If this sounds familiar, try a simple test. Take a document and conduct a "language audit." With a highlighter, mark every instance of words such as result, benchmark, indicator, goal, milestone, etc. Then go back to see if there is any pattern in the use of these terms. If there is no pattern, then consider the language convention offered here as a reference point. See if you can match the terms in your document to the definitions offered here, and use this relationship to create a common language.

The following definitions provide the conceptual starting point for results-based decision-making and budgeting:

> Result or Outcome:[4] A "result" is a bottom-line condition of well-being for children, families, or communities. Results are matters of common sense, above and beyond the jargon of bureaucracy. They are about the fundamental desires of citizens and the fundamental purposes of government. The results we are discussing are not "owned" by any single government agency or system. By definition, they cross over agency and program lines. Results include children born healthy, children ready for school,

[4]In some parts of the country the term "outcomes" has taken on a political meaning very different from the way in which we use the term here. We therefore use "outcomes" and "results" interchangeably to describe conditions of well-being for children, families, and communities. Such statements of well-being often span conventional political boundaries and provide a common ground for those with widely different ideas about how best to achieve those outcomes. This use of the term "outcomes" stands in contrast to its use in debates about outcomes-based education, where it is used to describe new approaches to measuring a student's knowledge and skills.

children succeeding in school, young people avoiding trouble,[5] stable and self-sufficient families, and safe and supportive communities. These are outcomes that we want for our own families, children, and communities. If we define results carefully, they will still be important in 10, 50, or 100 years. And because they have that kind of staying power, they are the right place to start thinking about what we want to achieve, and how we can get there from here.

Indicator: An "indicator" is a measure, for which we have data, that helps quantify the achievement of a desired result. Indicators help answer the question: "How would we know a result if we achieved it?" A result is not directly measurable by any single piece of data. There is no one complete measure of children succeeding in school or staying out of trouble. Examples of indicators are: rates of full immunization for children ready to start school, reading and math achievement scores, high school graduation rates, and rates of teen pregnancy and drug use. An essential element of this definition is that the data for an indicator are currently available. This is not about what we wish we knew, but about real-world information actually produced.[6] And as our data systems get better, we can add to the list of indicators.

Performance measure: A "performance measure" is a measure of the effectiveness of agency or program service delivery. These are measures of how well public or private agencies and programs are working. Typical performance measures address matters of timeliness, cost-effectiveness, and compliance with standards, such as child abuse investigations completed within 24 hours of a report, or the cost of child support collections for each dollar collected. Such measures are absolutely essential to running programs well. But they are very different from results and indicators. They have to do with our service response to social problems, not the conditions we are trying to improve. It is possible, sometimes common, for individual programs to be considered successful, even while overall conditions get worse.

---

[5]Most of the results given as examples in this paper are stated in positive, not negative, terms, with this one exception. It is, however, a common refrain to hear people say that they want their kids "to stay out of trouble." Someone recently offered "children with hope" as a positive version of this result. It's not quite as sharp, but it may be better.

[6]Note that, unlike the positive nature of result statements, indicators are almost always based on negative data. The reason is simple: Most of the data we actually collect is for things that go wrong.

The most important distinction in this set of definitions is between ends and means. Results and indicators have to do with ends. Performance measures and the programs they describe have to do with means. The end we seek is not "better service" but better results. The distinctions will help us describe budgeting processes that are built on clear thinking about what we wish to achieve and the strategies we choose to get there.

B. Turning the Curve: Defining Success in a Complex Environment

We often set ourselves up for failure in our work on family and children's well-being by creating unrealistic expectations and impossible standards for success. A large part of this problem is attributable to the way in which we use data to define success or failure. The typical approach to defining success is what we call, for want of a better term, "point-to-point" improvement. If the juvenile violent crime arrest rate is now 506 per 100,000 youths, we tend to define success as reducing this rate to 450 over the next two years.[7] This kind of definition of success is a setup. Most social conditions are more complex than this. These conditions have direction and inertia. This is reflected in a baseline, which is more often than not headed in the wrong direction. These directions can very rarely be changed quickly. Sometimes the best we can do is to slow the rate at which things get worse before we can turn the curve in the right direction. This is a more realistic way of thinking about success (and failure). Success is turning away from the curve or beating the baseline, not turning on a dime to achieve some arbitrary lower target.

There are at least two kinds of baselines that should serve as reference points for our (self) evaluation of success and failure in a result framework.

> Indicator baseline: The first is the baseline for each of the indicators. This baseline shows us where we have been on such measures as low-birthweight, or teen pregnancy, and where we are headed if we continue on our current course. These baselines can be used to show expected changes due to demographic or economic changes, such as the predicted increase in juvenile crime due to the coming growth in the population aged

[7] "Kids Count Data Book 1996: State Profiles of Child Well-Being," The Annie E. Casey Foundation, Baltimore, Maryland, page 21.

14 to 24. [8] (and they can help establish realistic targets for future performance, as discussed in Section IV.)

Cost of Bad Results (COBR) baseline: The second baseline is the companion-cost baseline. In this case the cost we need to consider is the "cost of bad results." Much, if not most, government spending for children and families, other than elementary and secondary education, is spent to address bad results: children born unhealthy, children not ready for school, not succeeding in school, not staying out of trouble. The costs of these bad results show up in both governmental and non-governmental expenditures. It is possible to measure and track these expenditures, and to begin to frame our social and fiscal policies in terms of reducing the growth in these costs.

Each baseline, in turn, has two components: an historical component and a forecast component. Forecasting is at best an inexact science, and forecasts should reflect a reasonable range of possible future courses – high, medium, and low, or optimistic, best guess, and pessimistic. As Yogi Berra once said, "Forecasting is difficult, especially about the future."

While forecasting can be difficult and even risky, the forecast component is very important. First, it communicates a powerful message about what we can expect to happen if we stay on our current course, and it can be used to frame the fundamental question in this work about whether that expected course is an acceptable one. Second, it provides a reference against which to look at data as it all comes in and make judgments about how we are doing month to month, quarter to quarter, and year to year. These kinds of processes can and should be dynamic,[9] using data to test ourselves and our strategies on a regular basis.[10]

Two other uses are worth mentioning, although we will not discuss them much here. The cost-of-bad results baseline can help set up a different way of approaching how we finance the investments necessary to turn the curve. When COBR analyses are completed, they are certain

---

[8]See "Catching a Coming Crime Wave," Scientific American, June 1996.

[9]What if the media anxiously awaited the release of monthly immunization data, as they now wait for the latest unemployment statistics? It may be that we need to think about creating a Bureau of Children's Statistics that could do for this area what the Bureau of Labor Statistics did for the labor movement.

[10]There is a growing literature on self-evaluation. See "Improving Evaluability Through Self-Evaluation," Charles L. Usher, Evaluation Practice, vol. 16, no. 1, 1995, pp. 59–68, or "Empowerment evaluation: Knowledge and tools for self-assessment and accountability," D. M. Fetterman, S. Kaftarian, and A. Wandersman, Thousand Oaks, CA: Sage, 1995.

to show the high cost of bad results and the relatively meager amounts embedded in the total cost now devoted to turning the curve. This picture is a first step in discussing the tangible financial benefits of an effective investment strategy to turn the cost curve, and may open the door to some non-traditional ways to finance that investment.

The second use will be controversial with the research community, but a well-established baseline is a kind of substitute for a control group in very complex environments. If we can show that our success at turning the curve(s) had some timely relationship to a set of strategies at scale (and that we were not just the beneficiaries of a fortuitous change in economic or demographic conditions), then we have credible, **circumstantial** evidence that these efforts are paying off. We will never be able to answer the cause-and-effect questions at the systemic level in the way we would like, but baselines, and our performance against baselines, can be a powerful, if still not fully satisfying, substitute.

Baselines are therefore an essential component of results-based decision-making and budgeting. Without baselines, we are blinded to the reality of complex problems and complex spending patterns. We are limited by systems that inaccurately measure progress and that skew decision-making away from preventive investments. Baselines allow us to think about problems in multi-year terms and to avoid the oversimplifications that accompany year-to-year or point-to-point comparisons.

Results-based budgeting uses baselines as the starting point for serious decision-making. The purpose of results-based budgeting can be reduced, in its simplest terms, to finding effective ways to improve our performance against the indicator and cost baselines for the most important results for children, families, and communities.

C. What is a Strategy Map (Anyway)?:

This paper presents a strategy map for implementing results-based budgeting. What is a strategy map?

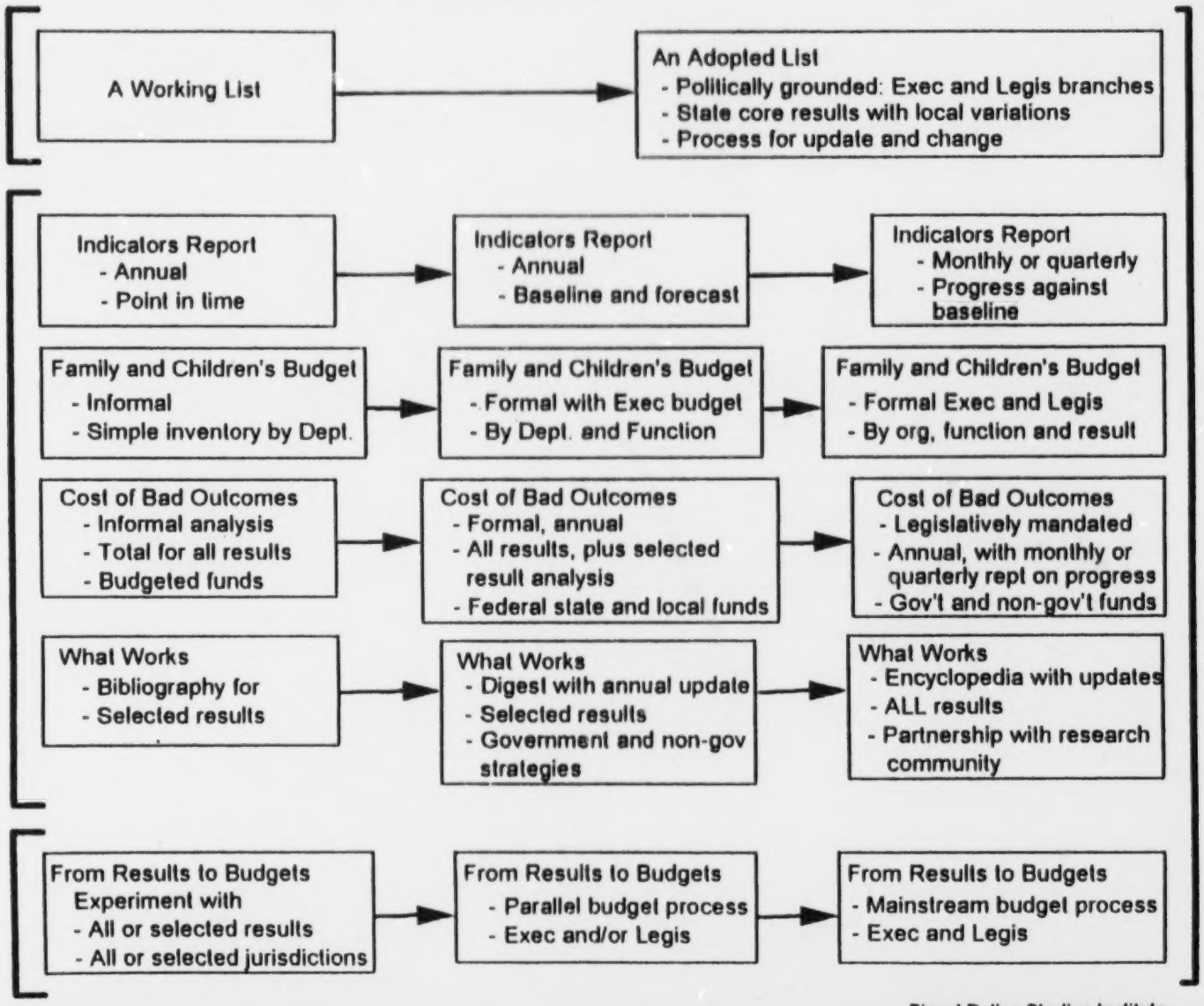
A strategy map is a format for displaying the complete implementation of a complex effort over time. Each box in the strategy map represents a "functional plateau," that is, a level of competency or functioning along a given developmental track. The map is organized by tracks and subtracks, and shows the progression of growing competency along each track. Each functional level is described in terms of what can be done at that level, or what products or functions characterize performance at that level.

What is not shown is as important as what is shown. Not shown are the actions required to get from one level to another. Usually this is the hardest part of the work, and the part that takes up the most space in conventional work plans or progress documents. In effect, a strategy map is a results-based document that derives a progression of concrete accomplishments from a defined end state.

A strategy map can be used as a rough work plan, and has been used this way in a few places. Strategy map "technology" can also be applied to other complex undertakings in order to help frame how a complex, multi-year undertaking can and should unfold. It has been used, for example, to show the structure of a total reform effort for family and children's services in several states. One great strength of this type of format is its power to communicate to a wide variety of audiences.

RESULTS BASED BUDGETING

A Strategy Map


Results and Indicators
A Working List
An Adopted List
- Politically grounded: Exec and Legis branches
- State core results with local variations
- Process for update and change
Decision Making Tools
Indicators Report
- Annual
- Point in time
Indicators Report
- Annual
- Baseline and forecast
Indicators Report
- Monthly or quarterly
- Progress against baseline
Family and Children's Budget
- Informal
- Simple inventory by Dept.
Family and Children's Budget
- Formal with Exec budget
- By Dept. and Function
Family and Children's Budget
- Formal Exec and Legis
- By org, function and result
Cost of Bad Outcomes
- Informal analysis
- Total for all results
- Budgeted funds
Cost of Bad Outcomes
- Formal, annual
- All results, plus selected result analysis
- Federal state and local funds
Cost of Bad Outcomes
- Legislatively mandated
- Annual, with monthly or quarterly rept on progress
- Gov't and non-gov't funds
What Works
- Bibliography for
- Selected results
What Works
- Digest with annual update
- Selected results
- Government and non-gov strategies
What Works
- Encyclopedia with updates
- ALL results
- Partnership with research community
Decision Making Process
From Results to Budgets
Experiment with
- All or selected results
- All or selected jurisdictions
From Results to Budgets
- Parallel budget process
- Exec and/or Legis
From Results to Budgets
- Mainstream budget process
- Exec and Legis

THE FINANCE PROJECT

Fiscal Policy Studies Institute

III. A STRATEGY MAP FOR RESULTS-BASED BUDGETING

This section describes each element of the results-based budgeting strategy map shown on the preceding page, and presents ideas about how to put these elements in place. The strategy map includes three main tracks: Results and Indicators, Decision-Making Tools, and Decision-Making Process. The basic progression here is simple. Creating a set of results and indicators (track 1) lays the groundwork for developing new decision-making tools (track 2), which informs, and to some extent makes possible, a new kind of decision-making process (track 3). The purpose of this effort is to make decisions that lead to improved results for children, families, and communities.

Track 1: Results and Indicators:

Since results and indicators provide the starting point for much of the other work, development of a results-and-indicators framework must come first. This does not mean that it is necessary to complete track one before going on to tracks two and three. The work on the three main tracks will necessarily proceed in parallel sequence (See the discussion of sequencing in Section IV.)

Level 1: A Working List: The beginning stage of this work usually takes the form of a working list of results and indicators developed by a collaborative group charged with identifying desired conditions of well-being for children and families. It is possible to get started without a formal, politically sanctioned process incorporating broad participation. But the strength of this beginning work depends in part on the capabilities of the group and its political legitimacy and credibility.[11] More political standing at the start will allow the work to progress faster and give it a better chance to take root. But less formal starting points are possible and not uncommon.

There are now many examples of results-and-indicators lists that have been developed around the country and that can be used as reference points in states and localities just beginning this work. Among the best known are Oregon's "benchmarks" and

[11]See "Toward New Forms of Local Governance: A Progress Report from the Field," Center for the Study of Social Policy, January 1996.

Minnesota's "milestones." Georgia's Policy Council for Children and Families recently adopted a framework of "results and benchmarks" for families and children, shown on the next page.

GEORGIA'S RESULTS AND CORE BENCHMARKS FOR CHILDREN AND FAMILIES

IMPROVED CHILD HEALTH: HEALTHY CHILDREN

- Percentage of infants born (1) weighing 2500 grams (5. 5 pounds) or more, (2) to mothers receiving prenatal care in the first trimester, and (3) to mothers who did not smoke or drink during pregnancy.
- Percentage of children appropriately immunized by age two.
- Pregnancy rate among school-age girls.
- Percentage of children having untreated vision, hearing, and health problems at school entry.

IMPROVED CHILD DEVELOPMENT: CHILDREN READY FOR SCHOOL

- Percentage of low-income students in Head Start or pre-kindergarten programs.
- Percentage of kindergarten students in pre-school or child care programs.
- Percentage of kindergarten students passing the Georgia Kindergarten Assessment Program.
- Percentage of students who are two or more years overage in the third grade.

IMPROVED SCHOOL SUCCESS: CHILDREN SUCCEEDING IN SCHOOL

- Percentage of students who are absent 10 or more days from school.
- Percentage of students performing above state standards on curriculum-based tests at grades 5 and 11.
- Percentage of students scoring above national median on normal achievement tests at grade 8.
- Percentage of students who graduate from high school on time.

IMPROVED FAMILY FUNCTIONING: STRONG FAMILIES

- Percentage of new families with (1) mothers who completed high school, (2) mothers who are age 20 or older, and (3) the father's name recorded on the child's birth certificate.
- Percentage of teenage mothers who give birth to another child.
- Rate of confirmed incidence of child abuse or neglect.
- Percentage of children in foster care who are placed in a permanent home.
- Percentage of youths arrested.

IMPROVED ECONOMIC CAPACITY: ECONOMICALLY SELF-SUFFICIENT FAMILIES

- Percentage of children living in poverty.
- Percentage of female-headed families with children living in poverty.
- Percentage of AFDC recipients who leave public assistance because of employment or higher incomes.
- Rate of growth in employment.
- Rate of unemployment.

One of the most important challenges in this stage of the work is to keep the list of both results and indicators short. It is not hard to come up with a long list of results. And any public service professional can easily list 20 or more indicators for a given result. It is harder to be spare. For each result, the group should identify three or four primary indicators to represent what this result "means for us" in measurable and measured terms. This is difficult, but absolutely essential to the clarity and coherence of the later work.

Indicators that are not selected here are not discarded, but can be placed on a second list for use in later parts of the process. It is also advisable to keep a third list of desired indicators where data needs to be developed or improved. Over time, groups may add to or move indicators from one list to another within this structure.

The following criteria can be used to select the "primary" indicators from a longer list of candidates.

Communication Power: One of the principal purposes of a result/indicator framework is to communicate with the public and other constituencies about "how we're doing" on child, family, and community well-being. It is possible to think of this in terms of a public square test. If it were necessary to stand in a public square and explain, with only two or three pieces of data, "what we mean in this community by children succeeding in school," what two or three pieces of data would you use? Obviously, you could bring a thick report to the square and begin a long recitation, but the crowd would thin out quickly. No one will listen to, absorb, or understand more than a few pieces of descriptive data. They must be powerful, common sense, and compelling, not arcane and bureaucratic measures. The point here is to achieve power and clarity with diverse audiences.

Proxy Power: Another simple truth about indicators is that they run in herds. If one is going in the right direction, chances are that the rest are as well. You do not need 20 indicators telling you the same thing. Pick the indicators that

have the greatest proxy power, i e., that are most likely to match the direction of the other indicators.

The second important dimension of the "proxy" criteria is the extent to which there is an established relationship between the data element and the result it represents. Is the indicator a good "proxy," not just for other indicators but for the result itself? The strongest linkage is, of course, based in research. But less formal common-sense linkages can also be important and useful.

Data Power: Last, but not least, it is important that the indicators we choose to "represent" the result are ones for which we have high-quality data and that allow us to see progress or the lack thereof on a regular and frequent basis. Ideally, such data elements would be updated on a monthly or quarterly basis, with a relatively short time lag. This would enable those overseeing this process to plot the new point on the curve and assess how we are doing in relation to the baseline on a regular basis. In a recent conversation in one state, we identified the youth health-risk survey data on teen suicide as a good indicator of the result of "children avoiding high-risk behavior." The indicator met two of the criteria, but not the third, since the survey was only conducted every two years. It was therefore considered reasonable to place this indicator on the "secondary" list.

In two recent local efforts, one in the Lamoille Valley region of Vermont, the process of choosing primary and secondary indicators involved ranking each of the candidate indicators on these three criteria. A high, medium, or low rank was assigned for each indicator on each criteria. For the data- power ranking, the participants also assessed the availability of both current and historical baseline data. The expense of collecting the data was coded with one-, two- or three dollar signs. This process helped build consensus on a list of primary indicators that could be ranked high in all three categories, with current and historical data available at reasonable effort and expense.

It is important to view the list of results and indicators as a functioning whole in order to be sure that the statements are internally consistent and complete. It may also be valuable to consider the idea of "checks and balances" between selected indicators. It should be possible to examine the extent to which improvement on one indicator may

be made at the expense of another. By testing an indicator set for natural checks and balances, and defining success as improvement on complementary indicators, we improve the chances of real progress. Consider as examples the relationship between rates of children in out-of-home care and injury/death rates from child abuse, or rates of welfare dependency and rates of poverty. In each case, having both indicators on the list can serve as a (partial) safeguard against the later development of narrow or expedient strategies that solve one problem by creating another. This safeguard structure does not have to appear in (and expand) the primary list of indicators. It is also possible to link primary indicators with complementary indicators on the secondary list.

Level 2: A Politically Grounded List: If results are to serve as a basis for serious decision-making, then they must attract or acquire political standing. That standing should be grounded in both the executive and legislative branches. In addition, a fully developed state framework should support the development of local results and indicators and provide for local variations that reflect local needs and priorities. This should also include a process for periodic review, update, and change at both the state and local levels. To date, only one state, Oregon, has developed a results/indicators framework that substantively meets all of these criteria. The state of Oregon and a number of counties in Oregon have grounded their use of benchmarks in state and local law. A number of states, cities, and counties are also making progress in this direction. The approach taken to this work by Oregon and Georgia is discussed in the boxed sections.

In contrast, at least one state has pursued the results/indicators framework as an invention solely of the executive branch, and its utility has been seriously eroded by contention over ownership. It is, of course, impossible to structure such a process without political controversy. But there are many precedents for bipartisan work on children's issues, and executive-legislative and state-local partnerships on results are within political reach.

There are two good approaches available to link state and local development of results and indicators. States can create a core list to which counties and communities can add (Georgia's approach), or they can create a comprehensive list from which counties and communities can choose, add, and modify (Oregon's approach). Either approach

works. One approach that does not work is for a state to develop and mandate a list that everyone must use without variation. State processes must treat local partners with respect, and must honor legitimate variations in local needs and priorities.

A Politically Grounded Results Framework

How Oregon Did It:

Oregon appears to be the only state that has established a results/indicators framework that is politically grounded in both the executive and legislative branches.

Oregon's framework includes approximately 260 "benchmarks" grouped within approximately 17 "strategic goals " Oregon's use of the term "benchmark" is closely aligned with the definition given above for "indicator." And Oregon's use of "goal" or "strategic goal" roughly corresponds to the above definition of "result." The benchmarks cover all aspects of quality of life for the state and its citizens, going well beyond matters of child and family well-being to include environmental, economic, and other measures.

Oregon used a broadly based process to develop the benchmarks framework, beginning in 1988 and leading to the issuance of the first "Benchmarks Report" by the Oregon Progress Board in 1991. This long development process involving citizens from all parts of the state helped create a political foundation that made bipartisan legislative endorsement possible. Since their introduction, the benchmarks have been formally adopted by successive sessions of the state legislature.

Some of the best work with the benchmarks framework is being done by county governments in Oregon. At least seven of Oregon's 36 counties are voluntarily developing comprehensive, locally oriented benchmarks systems. And several counties, notably Multnomah and Marion, are using the benchmarks in their budget processes. Benchmarks are also being used by Oregon's 13 key industries in their planning processes, and by private organizations such as the Oregon Community Foundation and the Portland United Way.

A Politically Grounded Results Framework

How Georgia Did It:

Georgia offers another perspective on how to develop a politically grounded results-and-indicators list, or in Georgia's parlance, results and benchmarks. Georgia has a long history of innovation in children's services. Among other efforts, the Family Connection initiative, started in 1991, has brought community leaders together in 71 counties and communities to assess the needs of children and families, and to take action to improve the quality of their lives.

In 1994, the Governor's (interim) Policy Council for Children and Families issued a report entitled, "On Behalf of Our Children: A Framework for Improving Results." This report recommended that the state establish a framework for results accountability. The framework will include "community partnerships" as a next stage in the development of The Family Connection, and as a basis for local results accountability.

In March 1995, the legislature passed SB256, a broad measure that made permanent the Policy Council and set out an ambitious agenda for creating local "partnerships" that would form the basis for state and local work on improving the well-being of families and children.

In May 1995, the Interim Policy Council set out to create a results-and-benchmarks framework. The Council created a Task Force on Results Accountability, which consulted broadly with state and local leaders, and, in a series of working sessions, drafted a working list of results and benchmarks. The Task Force completed the development of this working list by July 1995. The list was presented at a statewide Family Connection conference in August 1995.

At the time, the Council made clear its intention to use this list as a "core" list, which would allow counties and communities to add results and benchmarks, as appropriate, to suit local needs and priorities. The Task Force sponsored broad-based community meetings, and structured an open process to gather comments on the proposed list of results and benchmarks. The Task Force adopted several changes in response to the suggestions received through this process. In November 1995, the Task Force issued its final report, which included the recommended results and benchmarks for children and families and a proposed process for building accountability for those results. In January 1996, the new Policy Council convened for the first time. And in March 1996, the Policy Council adopted the Task (continued)

BEST COPY AVAILABLE

Georgia (Continued)

Force recommendations, with some minor modifications, as the basis for its ongoing work.

In July 1996, the Policy Council selected 10 partnership communities to serve as the vanguard in creating a locally based system of accountability for results and benchmarks. The designated communities will have direct access to the Policy Council to negotiate increased flexibility for resources and policy changes, and will receive support and technical assistance for implementing results-based accountability systems, as well as strategies for improving the well-being of children and families.

The Policy Council has initiated the work necessary to support the new accountability framework, including the development of a benchmark report, with 5-to-10-year baselines, for the state and each of the 159 counties.

While the results and benchmarks list itself has not been adopted by the Georgia legislature, the process for developing it is grounded in state law. Georgia's approach is instructive on many counts. Perhaps most importantly, Georgia's approach established a solid results/indicators framework in about 10 months.

Track 2: Decision-Making Tools:

It can be argued that one of the reasons our leaders tend to make shortsighted budget decisions is that they have only "one-year-at-a-time" decision-making tools to work with. If we are going to make good 20-year decisions about the well-being of children, families, and communities, then we need 20-year decision-making tools. We literally need to "retool" our decision-making process. Following is what might be considered a minimum tool set for this kind of decision-making process.

2.1 Indicators Report: An indicators report is one of two pictures we need at the beginning of the decision making process. We have good examples of beginning-stage indicator reports that have been developed as part of the Annie E. Casey Foundation's Kids Count project, and in other work by state and local governments and advocacy organizations. Indicators reports can be viewed in terms of three stages of development:

Level 1. Annual Point in Time: Such reports generally begin as a report card showing the status of families and children in the state or locality on each indicator **at one or two points in time**, usually the most recent data compared to some prior period. We now have a wide range of good examples of how to do this. The Kids Count Reports provide a picture of national and state standings on 10 key **indicators** of child and family well-being. These are being prepared at the county and (major) city level in many states. An interesting note concerning the Kids Count report is that the indicators do not tie to a stated result structure. It is arguable that there is an implied relationship, and, in fact, there is a good link between the Kids Count indicators and the results lists developed in many state and local efforts. But Kids Count indicators could possibly be strengthened by some more direct link with stated results for families and children.

Level 2. Annual with Baseline and Forecast: The next stage of development for indicators reports provides not just point-in-time data, but a true baseline for each indicator. As noted above, forecasts must include some recognition that forecasting is an inexact art; high, medium, and low forecasts are a common way to address this difficulty.

Some reports attempt to shortcut the baseline picture by offering two points in time and comparing performance between these two points. A crude two-point trend is inferred. This kind of analysis can answer simple questions, like "Are things better or worse than 10 years ago?" Kids Count has used this kind of format, and it has proved a powerful communications tool. But point-to-point comparisons, particularly over such a long period as 10 years, can mask the real trends. You cannot tell from such analyses, for example, whether the problem peaked in that 10 year period and is now declining or whether it bottomed out and is now getting better, or some other more complex picture. In fact, these last questions are the most important for policy purposes. So we need to move beyond reports with two-point comparisons, to a next-stage of development with real baselines, including forecasts. The 1996 Kids Count report moves in this direction by including 19-year historical baselines at the national level in the pocket guide accompanying the main report.

An Indicators Report with Baselines

How Vermont Did It:

In January 1996, the Vermont Agency of Human Services issued the **fourth** annual edition of its report on "The Social Health Status of Vermonters." The report shows Vermont's performance on 42 indicators in 8 categories of citizen well-being (Families and Children, Teens, Public Health, Economic Security, Safety/Crime/Corrections, Education, Elderly, and People with Physical or Mental Disabilities).

The report, one of the best in the country, is clear and easy to understand. And it is one of the only reports to present trend lines for data on family and child well-being. The report presents a 10-year or longer trend picture for more than half of the indicators (27 of 42). In 1995, the Agency of Human Services (AHS) and the Department of Education, acting in partnership, issued a set of Community Profiles that provided information on education and social health status for every school district in Vermont. These reports showed the most recent performance of the school district (or county, where school district information was not available) in comparison to state data. For many indicators, three-to five- year trends are presented.

The display of AHS data by school district represented a real breakthrough, both in terms of providing useful local analysis and in the work to link the efforts of the education and health and human services systems. Both the state report and the Community Profiles have received wide press attention in Vermont and have been used in many parts of the state to support community-planning efforts to improve results for families and children.

In future editions, the two departments are considering adding a forecast component to the state-trend data, and full alignment of the data analysis with the family and child results framework now being developed.

Level 3. Progress Against Baseline Report: Arguably the last stage of development on the indicators report track is a **structure** that provides monthly or quarterly reporting against the baseline forecasts developed in level 2. This structure provides the basis for the "continuous improvement" feedback loop that is central to the ultimate success of results-based systems. Comparison of actual data to baselines allows us to judge whether the strategies we have adopted to turn the curve are working or not. There are few, if any, examples of this type of report in public human services at this time.

2.2 Family and Children's Budget: There are three places in the country that have produced a children's budget on a regular basis. The longest-running children's budget is produced by Los Angeles County, which has issued this analysis every year since 1986. The initial work was done by the Los Angeles Roundtable for Children (tracking changes between 1980 and 1984) and is now supported by the Children's Planning Council and the Chief Administrative Office's budget section. Oklahoma has produced a children's budget since 1990. And Kansas has produced and published a children's budget as part of the Governor's budget submittal since 1993. There are numerous examples of children's budgets produced for one or two years, or produced by an organization outside government. States in this category include California[12] Georgia, Illinois, Maryland, and New York.[13]

The principal complaint about children's budgets is that they require a lot of work and are not used much in the budget process. We believe that the main cause of these sometimes legitimate complaints is that children's budgets are not conceived as part of a larger decision-making framework in which they have a defined role. They almost never get beyond the Line-Item Inventory stage discussed below, and never get to be more useful tools. These are, of course, connected, and somewhat circular, problems. Why spend the time developing a more sophisticated version of a children's budget if the current version is not used?

---

[12]California Children's Budget Data Report 1996–1997, R. C. Fellmeth, J. D., S. Kalemkiarian, J. D., and R. Reiter, Ph.D., Children's Advocacy Institute, 5998 Alcala Park, San Diego, CA 92110.

[13]We welcome corrections or additions to this list.

There are arguably three stages to the development of a family and children's budget:

Level 1. Line-Item Inventory: This is an aggregation of the program line items associated with spending for children as they are represented in the current operating budget. Such budgets have limited, but sometimes important, uses in assessing changes in total spending for children. An analysis that compares growth rates in state revenue with growth rates for children's programs can answer simple questions about the children's share of state resources and can be completed readily in states and localities with children's budgets.[14]

Level 1 budgets sometimes start as informal, behind-the-scenes summaries. In one state where this type of analysis was prepared, the release of the children's budget was delayed due to concern that its premature release could produce a political backlash against children's programs. It was thought that children's programs could be seen as consuming too great a share of state resources in a time of budget cutbacks. In other places, such as Los Angeles, the fact that children's programs made up a much larger proportion (one-third) of the county budget than generally thought, sparked greater attention to the needs of families and children on the part of policy-makers and business leaders. This highlights the essential political nature of these processes and the need for the tools discussed in this section to be considered carefully as part of a larger plan, not taken as isolated documents to be produced on their own.

Level 2. By Function: A second-level budget goes beyond the simple aggregation of existing line-item spending and presents (and analyzes) spending across agency and categorical lines by function. Oklahoma's budget provides this kind of picture, showing spending for child care, mental health, and eight other functional categories. This kind of picture makes the budget somewhat more useful because it allows at least a preliminary assessment of the interrelationship between program expenditures and their combined

[14]As an example, an analysis completed in Kansas for the FY95 budget showed estimated state revenue growth of 4.4%, while proposed children's expenditures increased by only 1.2%.

effectiveness. The Los Angeles County children's budget also provides a higher-level summary, classifying expenditures into eight functional categories, including a category for prevention expenditures. Such functional distinctions can set the stage for efforts to improve coordination of service delivery. In Los Angeles, the classification of program expenditures by functional/service area helped advance coordination across county departments by showing areas of related investment and common interest. Functional classifications also become useful in creating an investment case for children's spending, as discussed in the "cost-of-bad-results" section below.[15]

Level 3. Results-based: The "final" stage of children's budget development builds on the previous stages and provides not just line-item and functional pictures, but a results view of expenditures as well. We are just beginning to understand what a results-based budget document looks like. Perhaps the most advanced results-based budget in current use is the Multnomah County, Oregon, budget, which shows the relationship of the County agencies to the County's urgent benchmarks, and provides a summary of both current and new efforts to address each urgent benchmark. A similar effort is under way at the city level in Pasadena, California. It is likely that we will see more examples of results-based budget formats over the next several years.

The importance of the legislature in developing children's budgets specifically, and results-based budgeting in general, cannot be overstated. Legislators, like executive branch leaders, need to think about long-term accountability for child and family well-being. In addition, legislation provides the authority and stability for new budgeting approaches to be tried and adopted. A solely executive-branch approach often lasts no longer than the next election. With regard to children's budgets, there are two important roles for the legislature: First, the legislature should use the children's budget

[15] As discussed in the paper "Trading Outcome Accountability for Fund Flexibility," we may be able to make better sense out of state-local fiscal relationships if we begin to think about this funding in terms of "natural clusters" of funding. Natural clusters can be either functional clusters (e.g., all spending for child care, job training, etc.) or clusters that link prevention and remediation expenditures for a given population (e.g., spending for out-of-home-care and prevention of out-of-home care). Level 2 children's budgets give us a beginning picture of functional clusters that in turn provide a starting point for discussions about using dollars within clusters more efficiently.

in its budget deliberations and should amend and reissue the children's budget analysis following legislative action on the budget. Second, the production of a children's budget by the executive branch should be **required by law**, so that the budget is produced at the same time and with the same quality as the regular budget, and so that its production is not subject to the vagaries of executive-branch commitment and capacity.

Perhaps the most important argument for the formalization of children's budgets has to do with how they are used within the executive branch. If the executive-branch and the budget staff must prepare a children's budget each year and issue it as part of the formal budget submission, it is a good bet that people will ask, during the budget preparation process, "how will we look" when this document is released. For this reason, **legislatively mandated, executive-branch-produced** children's budgets may prove to be one of the best advocacy tools for children.

2.3 Cost of Bad Results: The idea of costing "bad results" starts with the idea of "good" results, the conditions of well-being we hope to achieve for children, families, and communities. From a fiscal perspective, much, if not most, government spending for children and families, other than elementary and secondary education, is spent in response to bad results: children born unhealthy, children **not** ready for school, **not** succeeding in school, **not** staying out of trouble. The costs of these bad results show up in both governmental and non-governmental expenditures. It is possible to measure and track these expenditures, and to begin to frame our social and fiscal policies in terms of reducing these costs. Such analysis will certainly show the enormous financial stake we have in improving results, and may set the stage for discussing alternatives.

If government were a business, we would be tracking the money spent on repair. If repair costs began to consume unreasonable amounts of corporate resources, we would do something about it. The first step would be to know how much was spent, and track repair costs over time. We would use this tracking system to see whether our preventive efforts to control repair costs were working. Our spending on bad results is (roughly) equivalent to business repair costs. We know that preventive maintenance is less expensive than repair. We have good reason to believe (though hard evidence is harder to come by) that preventing children's problems is less expensive than

remediating problems later.[16] The only way to reduce the high cost of bad results, over the long term, is prevention.[17] And the starting point for controlling these costs is to know what they are.

The question to be answered in this work is, "What costs exist today because we are not getting the results we want?" To put it another way, "What costs would go away if we achieved good results for all children?" When the question is posed this way, whole programs become part of the "cost-of-bad-results" answer. Welfare, foster care, juvenile justice programs, and all their attendant costs are part of the cost of bad results. It is possible to create a multi-year picture of all spending for children and families through a family and children's budget, and then cull out those pieces that fit the answer to this question. The total costs must then be adjusted for inflation and population growth so that we can see real changes over time. Production of this kind of analysis is closely linked to development of a children's budget, as discussed above. Children's budgets can provide the raw material for cost-of-bad-results analyses, and should be structured so as to anticipate the information needed for such analyses.

There are two principal uses for this work: **First,** as a tool to measure our success or failure in strictly fiscal terms. We need to track whether the strategies we develop to produce desired results are working to slow the growth and eventually to reduce the cost of bad results. **Second,** we can begin to think about the long-term financial benefits of improving the well-being of children and families and communities in real-dollar terms. This could lead to new ways to think about financing the investments in prevention necessary to make this happen. We can then ask an essential second question: "What expenditures are embedded in the total cost of bad results that are now devoted to turning the bad-result-curve?" The answer to this question will help identify the elements of an agenda for children and families that could turn the cost-of-bad-results curve.

---

[16]See the studies of the Perry Preschool, Women, Infants, and Children nutritional program, or Head Start as examples.

[17]It is important to note that not all prevention efforts are created equal. Poorly conceived or poorly delivered preventive programs may not be less expensive than repair.

There are three stages to the development of cost-of-bad-result decision tools:

Level 1: Total for All Bad Results: Work on the cost of bad results begins with the total costs for all results for a few simple reasons. Total program expenditures are known facts. We do not have to break programs up and split their costs between results. (We avoid, for the moment anyway, trying to answer questions like, "What portion of welfare expenditures is caused by teen pregnancy vs. children not ready for school?"). Second, when success is defined in terms of total cost of bad results, then cost-shifting "solutions," and moving costs from one system to another, or from the present to the future, are seen for what they are: irresponsible and ineffective. As with the indicators baselines, the cost-of-bad-results analysis should have both an historical and forecast component. Together, these provide a complete fiscal baseline for our work on "repair" costs.

At the first level of development, the analysis covers only budgeted funds for the level of government conducting the analysis, and may be an informal, behind-the-scenes analysis. To the extent possible, the analysis should identify those expenditures embedded in the total cost, which are now devoted to turning the curve. These expenditures usually make up only a small portion of the total expenditure on bad results, and represent the natural investment agenda for turn-the-curve strategies. (See the discussion of the "COBR Prevention Trap" in section IV.)

Level 2: Total plus Selected Bad Results: At the next level of development, the analysis adds, to the baseline and forecast for all results, the selective analysis of the costs of individual bad results. The principal purpose of individual results analysis is support for the decision-making process discussed below. In particular, the cost of individual bad results can help the team working on that particular result assess the potential fiscal benefits of turning the indicator and cost curves. Such analysis can be used to "sell" investments in a recommended turn-the-curve strategy. And in the best circumstances, such analysis can be used to craft a return-on-investment approach to finance some of the investment.

This subordinate analysis is actually harder to complete than the total analysis. It requires the development of some credible connection between a bad result and a specific cost. And given the complex interrelationship between programs and their effects, this may not always be possible. There is, however, some base of experience with this type of analysis. Individual cost of bad results analyses (or "cost of failure" analyses, as they are sometimes called) have been completed for teen pregnancy and its cost-effects in welfare, foster care, and Medicaid expense, for example. But research – particularly longitudinal research necessary for establishing systemic effects – is much more limited. (The much-used, perhaps over-used, Perry Preschool study is still among the best after two decades[18]). Computer models can be, and have been, used to build on what research is available and to establish a plausible cause-and-effect and cost-benefit relationship.[19]

The second-level development of cost-of-bad-results analysis should also go beyond a single level of government expenditure to look at total federal, state, and local expenditures. This becomes particularly important in looking at the cost-benefit effects of targeted prevention investments. In analyses of the welfare savings effects of employment and training programs, it is common to find that state general fund investments do not produce equal or greater general fund savings. But when federal savings are added (including, for example, 100% federal savings in food stamps benefits), the total cost/benefit relationship can become positive. Over the long term, it will be important to access savings that accrue to all levels of government from investments in improved results for children and families.

---

[18]See L. J. Schweinhart and D. B. Weikart, Young Children Grow Up: The Effects of the Perry Preschool Program on Youths Through Age 15, Ypsilanti, MI: High/Scope Education Research Foundation, 1980.

[19]The emerging literature on theory of change may also be relevant here. See New Approaches to Evaluating Community Initiatives: Concepts, Methods, and Contexts, edited by James P. Connell, Anne C. Kubisch, Lisbeth B. Schorr, and Carol H. Weiss, Roundtable on Comprehensive Community Initiatives for Children and Families, The Aspen Institute, Washington, D.C., 1995.

Level 3: All Results, All Fund Sources, with Progress Against Baseline: The most advanced form of a COBR analysis includes all preceding elements plus analysis of both governmental and non-governmental fund sources. Such an analysis would present baseline and forecast components for total public and private spending for children and families, and would provide for regular reporting against the baseline as part of the larger effort to track success on turning indicator and cost curves.

Cost of Bad Results

There is a long history of work on the cost of bad results on which to build. But none of the efforts to date have taken the approach put forward in this paper. The work generally breaks into two categories, those efforts that look at the total costs, and those that zero in on costs of specific bad results or conditions.

Total Cost Efforts: Perhaps the best-known of these analyses is the work from Austin, Texas, and more recently from Allegheny County, Pennsylvania, which compares the differential spending on social programs for low-income and middle-class neighborhoods. This analytic approach takes the credible assumption that such differences represent the potential savings that could be achieved if the problems of low-income neighborhoods could be successfully addressed. The results of this type of analysis can be dramatic, and can play a role in making the case for investments in prevention, and better use of service system resources. In addition, to this work, the Pew Foundation sponsored work on the "cost of failure" in the state competing for its system of care grants in 1994. Partial work on this analysis for Georgia is available.

Specific Result Efforts: Two types of result specific analysis have been done in the past. Research on the actual costs of certain conditions (e.g. teen pregnancy) through the various social systems (welfare, Medicaid, etc.) have been completed. And computer models have been created that simulate the costs of alternate cohort pathways through the service system and the cost-effects that may be imputed to certain investments that promote one pathway over another. The problems with this work are simple. The research base for making these connections is weak. Longitudinal studies are very expensive and in short supply. Our underinvestment in such studies sometimes leaves us with out-of-date information of limited applicability. The computer models and analyses that exist are built on a wide range of assumptions that also limit the utility of their results.

A national prototype analysis of the total cost of bad results is now under development with support from the Finance Project. This analysis breaks with the approaches discussed above and is built from the total federal, state, and local costs of total programs associated with bad results.

As discussed below, the total program cost approach avoids some of the credibility problems associated with splitting program costs by results. And this approach will allow the creation of multi-year baselines against which to measure success in turning the financial curves. This analysis should be completed in draft form by the end of 1996.

2.4 What Works: The Results to Budgets process calls for us to think systematically about what works to turn the cost and indicator curves. The answer to this type of question is not something that can be extracted from a review of the research literature. The reason, in part, is that we do not have, and never will have, all the research information we want or need. The answer can and should draw on lessons from the broadest base of experience, including research. But ultimately the answer for any state, county, or community involves people making their best judgments about what they think will work and what they are willing to stand up and defend in the public square.

But best judgment does not mean blind guesswork. There is a wealth of information on social experimentation that can be of tremendous value in this work (see the boxed discussion of "What Works"). Much of this work is pointing in the direction of comprehensive, cross-system, community-based strategies which combine effective governmental and non-governmental efforts (see the discussion of Tillamook County, Oregon, and MADD later in this section.)[20]The challenge is making this information accessible and relevant to the task of crafting strategies to improve results. The strategy map track for "what works" addresses the need to create this kind of useful and usable information.

Level 1: The work on this tool track can start as a bibliography for selected results. There are existing publications that summarize the documented effectiveness of programs. (See the sidebar discussion of employment and training, low-birthweight, and drug-prevention programs.) Identifying, assembling, and making these summary documents readily available should be a first step in supporting the work of decision-makers looking for effective strategies to turn the indicator and cost curves.

Level 2: Over time this could become a more structured form of support. A well-designed annual digest could be established with regular updates and information from both governmental and non-governmental sectors. The scope of coverage could grow to include a complete set of results and

[20] Cheryl D. Hayes, Elise Lipoff, and Anna E. Danegger, "Compendium of Comprehensive, Community-Based Initiatives: A Look at Costs, Benefits, and Financing Strategies." Washington, D.C.: The Finance Project, 1995.

BEST COPY AVAILABLE

indicators, cross-referenced to the results of states and localities with established frameworks.

Level 3: Eventually, this could become a national document or service, perhaps on line, with information related to all results. Such a service could provide a ready means of linking those who have operated or documented effective strategies, and those looking to replicate or adapt such strategies. It would even be possible to consider the development of an "expert system" that allows a user to interact with a specially designed computer program that helps navigate the complex literature to find the best answer to a particular problem.

At each stage of this work, this service can and should become a partnership with the research community to help get answers to the most pressing questions about the effectiveness, and in particular the cost-effectiveness, of strategies to improve child and family well-being. This means more spending on longitudinal studies, and more research involving cost-benefit assessments.

What Works

In January 1995, the Department of Labor published one of the best examples of a useful compendium of research findings on what works. What's Working (and What's Not): A Summary of Research on the Economic Impacts of Employment and Training Programs provides a comprehensive review of social science evidence on the economic impacts of employment training and education programs, including a review of over 100 research and evaluation studies. The report asks and answers tough questions about programs in four program categories (jobs for youth under 21, programs for disadvantaged adults, other sources of education and training, and re-employment programs for dislocated workers). The purpose is not to promote a particular approach (although the publication itself presumes some willingness to intervene in the labor market), but to give the lessons, both good and bad, from this work. A local task force, with a charge to turn the curve on family self-sufficiency, would find this a relevant tool.

This is, of course, only one of many examples of well-prepared "what works" tools. Others include the Packard Foundation publications The Future of Children, a series of reports designed to make information on effective programs accessible to decision-makers. The Spring 1995 issue, devoted to issues and strategies related to low birth-weight, would be of great value to a team working to turn the curve on healthy births. A recent publication, Making the Grade: A Guide to School Drug Prevention Programs, by Drug Strategies of Washington, D.C., reviews and rates a wide range of programs for preventing tobacco, alcohol, and drug use.

It is possible to imagine that similar publications could be systematically assembled, indexed, and made available to help connect decision-makers to what is known about what works to turn the indicator curves in a state or local results/indicators framework.

Track 3: Decision-Making Process:

Changing the decision-making process is the most sensitive, but also the most important, challenge in implementing results-based budgeting. It is not worth producing any of the decision-making tools discussed above, if they are not actually used to make better decisions about how to improve results. Without some commitment to reshape the decision-making process, the rest of this work is potentially wasted (see the discussion about sequencing in the issues section).

The strategy map suggests one possible sequence for developing a results-based approach to decision-making and budgeting. The logic is simple. Start with an experiment, allow the work to develop and prove itself over time, and eventually use results-based approaches as part of the mainstream decision-making and budgeting processes.

Level 1: An Experiment: Why experiment? First and foremost, it is impossible and imprudent to try to change the whole budget process at once over a single budget cycle. Second, these ideas are largely untested, and before committing to change we need to know if they can, in fact, work. Third, results-based approaches are more work (possibly a lot more work), and it is unwise, and for most government operations impossible, to staff this kind of undertaking without more resources. These resources must be earned. Fourth, political considerations make it necessary to move carefully, if not slowly. There are vested interests in the old budget approach. People already know how to play and win the annual budget game. A change of rules, even if it is geared to the long-term interests of children, will feel, and in some cases will actually be, threatening to some interests. Finally, the use of an experiment can increase the chances of early success by enlisting individuals (or organizations) that are most ready to innovate and change. Lessons from other change processes suggest that true innovators are rarer than those who eventually adopt and use new systems.

There are two ways to approach development of an experiment, which have to do with where the boundaries are drawn: The first approach involves a focus on a particular result or indicator. There are many cross-agency efforts already under way that provide natural forums for this kind of test. In one city in a midwestern state, a cross-agency task force was formed in 1995 with extensive business and community participation devoted to opening a jobs center for people on welfare. It is not too far a

BEST COPY AVAILABLE

stretch to imagine that this task force might broaden its charge to think about ways to "turn the curve" on self-sufficiency in the community. Within this framework, the jobs center could become a centerpiece of the community's "what works" strategy. This broader charge would allow various players to bring other resources to the table. (The business community could augment its support for center operations with jobs.) And the group might be able to impact indicators of family well-being that have been used to justify the center's development (like number of first entries onto welfare). It might be possible, in other words, to have more to show for their work than a new center. The center could be the first in a series of "what works" successes on the larger and more important challenge of family self-sufficiency.

The second way to draw boundaries around an experiment is to consider a broad array of results (up to and including the whole list) within a defined geographic area (county, city, school district, or neighborhood). There are a number of places that are considering testing this approach. Georgia's work with partnership communities will support the development of a results/benchmarks framework in selected sites and will help facilitate their experimentation with a results-based decision-making process. Other states are sponsoring partnerships between state government and local governance entities where results-based accountability is part of the work. A partial list of such states includes Iowa, Illinois, Maryland, Missouri, Nebraska, Oregon, Vermont, and Washington. These partnership sites represent natural forums for experimentation with results-based decision-making and budgeting approaches.

Level 2: A Parallel Process: If results-based decision-making and budgeting proves its worth at the experimental stage, it may become possible to operate a results-based decision-making process in parallel with the regular budget process for some longer period of time.

This does not mean we get rid of line-item budgets! Line-item budgets are here to stay, and for good reason. It will always be necessary to clearly identify the money that agencies and administrators have for agency and program operations, and to account for the legitimate use of these funds. While such line-item budgets may be necessary tools to operate government, they are not necessarily good tools to make decisions about what government does. Line-item budgets are aligned by government's highly categorical agency and program structure. They typically show expenditure

information for one or two years at best, with little or no information on use of resources across programs, agencies, or fiscal years.[21]

A results-based budgeting process can be used in parallel to develop (or at least inform the development) of the annual line-item annual budget process. It would be possible, for example, to use a results-based agenda as a reference point in thinking about how to move the system in one-year increments toward fulfillment of a longer-term plan for improving results. The results budget can then be cross-walked to the line-item operating budget.

Level 3: Mainstream Process: What does it mean for results-based budgeting to be the mainstream process for a state, city, or county? The answer must combine what we have seen implemented so far, as well as speculation on what is possible. Among existing budget processes, the most advanced results-based process is that used in Multnomah County, Oregon. While it does not have all of the elements discussed in this paper, the final budget document reflects past-year, current-year and next-year strategies to improve performance on urgent benchmarks. The county's benchmarking process provides tracking for progress on benchmarks. The individual agency budgets describe general linkages to the benchmarks and reference partnerships with other agencies, and include program-level performance measures with trend data.

The key feature of mainstream use of results budgeting, however, does not lie in the appearance of the budget documents, but in the way in which the content of those documents is developed, tracked, and used. A fully developed results budget process could employ a development process that is cross-categorical from start to finish, where the development of strategies and budget priorities emerges from cross-agency teams organized around each result. The result of these teams' work could then be translated by a "common-ground" team into a coherent and fundable multi-year agenda that crosswalks to individual agency budgets, and describes the interrelationship of government and non-governmental roles in turning the indicator curves. In the legislative branch, results could be used to structure committee assignments or to organize joint committee hearings by results. Agency budget line-item review would

[21]See the discussion in "From Outcomes to Budgets" (pp. 3–6) of the problems with current budget systems: shortsightedness, fragmentation, and process, not results, orientation.

continue, but become subordinate to this more strategic look at the use of resources to improve the well-being of children, families, and communities.

A Successful "Turn the Curve" Strategy

How Tillamook County, Oregon, Did It:

In 1990, the teen pregnancy rate in Tillamook County was 24 per 1,000 girls ages 10 to 17, worse than all but 5 of the state's 36 counties. Beginning that year, and continuing to the present, community leaders in Tillamook fashioned a community-wide strategy to change this condition. The strategy was simple: Get everyone – churches, public and private agencies, schools, health workers, and families to – acknowledge the problem and commit themselves to doing whatever they can to change it. The controversial nature of the challenge was actually turned into an asset. The widely different views of leaders and the institutions they represented helped motivate the community to get involved.

Between 1990 and 1994, the teen pregnancy rate decreased to 7.1 per 1,000 girls ages 10 to 17, the best rate in the state. Tillamook County does not attribute this success to any particular service, but rather to the combined effects of the community efforts. These included:

- Schools: added self-esteem and sexuality education to their curriculum.
- Churches: worked at opening up communication channels with teens, taught refusal skills, and promoted abstinence.
- County Health Department: with support from the County Commissioners, the department expanded clinic hours and changed policy to ensure that any teen who called the health department for information or services would be seen within 48 hours (not two to three weeks, as in previous practice).
- YMCA: sponsored a "teens at risk" program, providing recreation activities that kept teens busy and built up their self-esteem.
- Community College: worked with teens through the Tillamook Teen Parent Program to prevent second unintended pregnancies.
- Commission on Children and Families: funded teen-pregnancy prevention curriculum in the schools, as well as counseling and support groups.
- The Tillamook County General Hospital, with other partners, opened "Healthy Families of Tillamook County," a home visiting and parenting program for all newborns.

Other partners included the Women's Crisis Center, the Tillamook Family Counseling Center, the Tillamook Bay Child Care Center, the Tillamook Bay Community College, and others.

According to the Health Department summary, Tillamook County "found that forming partnerships and working together toward a desired result can bring about astounding results. Their turn-around was a evolutionary process, with new partners bringing contributions forward at different times." Given a catalyst and a target focus on a desired result, the same process can occur in other communities.

A Successful "Turn the Curve" Strategy

How MADD Did It:

This may seem like an odd example to include in a paper about government decision-making. But Mothers Against Drunk Driving (MADD) provides one of the best examples of people who set out to change a condition of well-being through a deliberate community-wide strategy of trying and testing things that work. And they have succeeded. We often look to the business sector for examples of how to make government work, and there is plenty to learn there. But MADD can teach us something different. They teach us not to wait for a federal grant, not to wait for the research community to tell us the answer, and not to measure our success by how many projects we have implemented or how much money we raised, but by whether we made a difference, and whether the trend line has slowed its growth, flattened, and begun to turn down. In this calculus of budgeting, numbers mean lives. MADD reminds us that we can change the rules of the game and win.

MADD was founded in California in 1980, and has grown to include hundreds of chapters in the United States and other countries. The work of MADD focuses on finding effective solutions for drunk driving and underage drinking, and supporting victims of drunk-driving crimes. Many of the actions that MADD has taken are familiar. These include direct actions such as Operation Prom/Graduation, the Red Ribbon campaign, designated-driver programs, court monitoring, and victim assistance programs; and support for federal, state, and local legislative changes, including age 21 drinking laws, license revocation and other penalties for repeat offenders, laws lowering the blood alcohol content limit for adults and setting "zero tolerance" for any blood alcohol content for those under 21, and victims' rights and compensation laws, among many other actions.

While MADD can not and does not claim full credit, the change in the curve is dramatic. After reaching a peak in 1980, the rate and number of alcohol traffic fatalities has steadily declined, from 25,165 in 1982 to 16,589 in 1994. What makes these statistics more important is the fact that there are approximately 60 alcohol-related **injuries** for every fatality. The **direct** cost of alcohol-related crashes is estimated at $44 billion in 1993. This estimate does not include pain, suffering, and lost quality of life, which raise the alcohol-related crash figure to $134 billion in 1993.[22]

A part from the impact on peoples' lives, the reduction in U.S. alcohol-related traffic deaths from 1982 to 1994 can be estimated to have saved $13.8 billion in **direct annual** costs.

Source: Publications and statistical summaries from Mothers Against Drunk Driving, Irving, Texas. Their cooperation and support is gratefully acknowledged

[22] Ted R. Miller, and Lawerence J. Blincoe, "Incidence and Cost of Alcohol-Involved Crashes," Accident Analysis & Prevention, vol. 26, n. 5, p. 583–591, 1994. Citation from MADD statistical summary.

IV. LESSONS AND ISSUES:

Even where results-based budgeting finds its way into the mainstream of budget practice, it will almost certainly continue to look and feel like a work in progress. If it is done well, results-based budgeting is a self-renewing and self-correcting process that will produce a continuing flow of lessons on what does and does not work, both for children and for budgeting. This section addresses some of the issues that have emerged in state and local work to date on results-based budgeting systems, and it summarizes what has been learned so far.

A. Turning the Curve in a Successful Environment:

Not all trend lines are going in the wrong direction. And results-based decision-making and budgeting is not solely applicable to systems or communities that must first stem the tide of worsening conditions. The education system and education indicators, for example, are showing signs of improving in many parts of the country. Health indicators for infants and young children have been improving nationwide for a number of years. (The Kids Count 1996 trend lines for infant and child deaths show steady declines from 1975 to 1993.)

The concept of turning the curve applies equally well to situations where the trend line is going in the **right** direction. The phrasing of the question becomes: "How can we turn the curve away from the baseline toward even better performance?" In Maryland during the mid-1980s, the trend line for children in foster care was already declining. The state's efforts to improve permanency for children in out-of-home care had the effect of creating a faster decline than the state and national pattern of caseload decline at the time. The state's investments "worked" by accelerating the rate of decline. In communities that are making progress in education, the results-based decision-making system can be used to ask how this success can be accelerated.

Equally important is the environment where there is a record of success to be sustained. It is not uncommon for systems to "rebound." In Tillamook County, the success of efforts to reduce teen pregnancy has created a new challenge to keep the pregnancy rate from turning back up. Because results and indicators have some staying power, results-based systems with regular reporting components could help to avoid shifting attention from successful areas and allowing problems to rebound. The system could highlight current performance in relation to the current baseline forecast and then quickly identify any rebound.

B. Targets vs. Baselines:

There is a natural temptation to establish specific performance targets for indicators, and several states have included out-year targets in their "benchmarking" documents. Many of these targets, however, bear no relation to baseline performance or to what is in fact achievable. Targets set without reference to baselines will at best detract from the credibility of the process, and at worst create false expectations and a perception of failure. Baselines are more difficult to produce. But they provide a more complete picture of current status in relation to the likely continued trajectory, and they provide a more solid base from which to describe alternative "success trajectories."

The use of targets makes more sense within a baseline structure, because the baseline helps people judge the likelihood of turning the trend line toward the target and then reaching it. Targets can be shown on baseline charts in one of two ways: either as an absolute standard (i.e., a flat line toward which we wish to drive the curve), or as a desired future path that turns away from the baseline in a specific way.

C. The Question of Sequence:

As noted above, the strategy map can be, and has been, used as a prototype work plan. It is possible to assign target dates to the performance levels along each track and use this to drive more detailed work plans. Clearly, the work on results and indicators must begin first. But after that, what is the best way to sequence this work? There are, at least, two possible approaches:

- Tools create a constituency for process experiments: One view is that the new decision-making tools will create a constituency for experimenting with or changing the decision-making process. If, for example, a set of trend-line charts could be created that show that the cost of bad results is growing faster than state revenue, this may spark interest in using a results-based process to turn the curves on conditions and cost.

- Process experiments create a need for tools: On the other hand, there is a risk that a lot of effort will go into creating tools for which there is no use (that is, the bookshelf trap). It may be best to create decision-process experiments and let these processes provide direction for creating results-based decision-making tools. A work group, for example, trying to craft a multi-year action plan and budget to turn the curve on juvenile crime will quickly need

baselines for indicators and costs, and resources to help in the "what works" part of the effort. The tools created by such work teams can become prototypes for tools that will be used in the larger process by other teams.

It is, of course, possible to think about moving on all three tracks at once. And this is in fact the most reasonable course. There is no need to wait for the results and indicators framework to be completed, politically grounded, and in place at the state and local level before proceeding with the other elements of work. Creation of some statewide tools like the children's budget and cost of bad results analysis at the total level can get under way while process experiments are commissioned that will develop tools for use on specific results and indicators. As with all else in results-based budgeting, the bottom line here is "what works" in the political environment. It is essential that the work be understood and supported by the state, county, or community leadership. And it may take some time to develop enough support for this work to proceed on a large scale.

D. Children's Budget "Sponsorship"

Should the children's budget be produced inside or outside of government? There are several reasons to hold a strong preference for producing the children's budget inside government, and specifically by the executive-branch budget office.

The first and most important reason is credibility. Credibility is a necessary, though not sufficient, condition for effective advocacy. But there is no reason to waste energy in budget discussions defending the numbers themselves. If they are produced by the official budget agency, then they will likely be accepted, and the real discussion about content and policy can take place.

Second, and perhaps more importantly, the fact that the executive branch must present a children's budget can have an important influence on the formulation of the budget. We know that advocacy for children inside government can be as important as the more visible advocacy outside of government. It is important to think about the way in which budgeting tools create opportunities for those with formal government responsibility to make good decisions for children and families. "Inside" production of such a document, of course, has a double edge, and can create a new political risk, a possibility that will not be lost on chief executives. This may make it harder to establish as a formal part of the process, or to retain, once established.

Finally, the "inside" production of the children's budget can strengthen the role of the advocacy community in several ways. Perhaps most importantly, the resources used for production can be put to other uses, including the development of analyses and communication tools that go beyond the budget summary itself. And it becomes possible to propose improvements in the children's budget presentation (such as the mandatory inclusion of an annual cost-of-bad-results analysis) without having to staff this work directly.

E. The COBR "Prevention" Trap:

Results-based budgeting tools and processes will serve to highlight the need for investments in early and broadly based prevention programs for children and families. As noted in the cost-of-bad-results discussion, this is as much a matter of good financial policy as it is a matter of conscience. We can no longer afford to keep treating problems after they occur. The commitment to prevention, however, is widely advocated but weakly practiced.

Given this predisposition to "advocate" for prevention, every professional likes to think of their program as contributing to prevention in some way. Even "deep end" services, like prisons, have program components devoted to rehabilitation, education, or job preparation that can be viewed as preventing recidivism and therefore crime. So when we pose questions in the cost-of-bad-results analysis in terms of "prevention and non-prevention," every program claims that it should be counted as a prevention program. The discussion can quickly descend into a useless debate about funding for "good" programs vs. "bad" programs. Having programs compete to be designated as prevention oriented completely misses the point of the cost-of-bad-results analysis, let alone results-based budgeting.

There is a simple solution to this problem. The cost-of-bad-results analysis is intended to identify the cost curve we wish to turn. It is possible to think of this analysis in two stages. First ask and answer the question: "What expenditures exist today because we are not getting the results we want?" Asked this way, the costs we must identify include whole programs. The Aid to Families with Dependent Children program, for example, existed, in its entirety because all families were not self-sufficient. The Medicaid program exists because people cannot afford or get access to health insurance. Another way to think about this first question is to consider what expenditures would disappear entirely if we achieved all good results. This total set of expenditures represents the cost curve we wish to turn.

The second question to be asked is: "What expenditures, embedded in this total, are now devoted to turning the cost curve?" This is the point at which we consider programs such as the immunization program within Medicaid devoted to reducing long-term costs of remediating health problems or the training and employment programs within the welfare system devoted to reducing the long-term costs of dependency.

We have, in essence, asked the prevention vs. non-prevention question in a way that does not stigmatize programs, or create false incentives to categorize expenditures one way or another. Remember that the first and last principle of results-based decision-making is "honesty with ourselves." It does us no good to do produce analyses, roll up expenditures, or sponsor decision processes that do not contribute to the actual process of turning the indicator and cost curves. The COBR analysis is, of course, only one test of this discipline. But because it sets out the cost curve we must turn, it is a crucial one.

F. The Various "What Works" Traps

One of the strengths of the approach outlined in this paper is the use of common-sense ideas and terminology to drive much more complex processes. The flip side of this approach, however, is the possibility that a commonly used phrase, like "what works," will be misinterpreted and misused. We use the phrase "what works" in two ways:

> First, what works to turn the indicator and cost curves? In this usage, we are looking for strategies, collections of programs, and approaches that span governmental and non-governmental sectors to impact the trend lines and turn them in the right direction.
>
> The second usage has to do with individual components of that strategy determining whether programs or elements of the overall strategy are working in one of two different ways.
>
> - Do they contribute to the overall "what works" strategy? Should they be part of the effort to turn the curve? Is there evidence that, in concert with other elements of the strategy, they contribute to turning the curve?
>
> - Are they working on their own terms? Are they well designed and well run? (See discussion of performance budgeting below.)

This latter question is the subject of performance measurement and accountability. Performance measures can be used to manage programs better, to get them to work properly. It is, however, possible to have a perfectly working program (performance measurement perspective) that contributes nothing to turning the indicator and cost curves. Among the best examples is the often-cited case of the school-based drug-education program that increased knowledge of drugs and the subsequent incidence of drug use among students receiving the training. No doubt the program could show that it increased knowledge of drugs (a performance measure), but it did not work in turning the curve on drug use (an indicator).

There are at least three ways that this phrase can be misinterpreted or misapplied:

The list of approved programs: If the state is assembling a list of what works and what does not work, then program staff and supporters will press to make sure that their program is on the "what works" list. In this view, the process leads to a list of approved programs that deserve more money, and a list of those that deserve less. This process is closely related to the prevention trap discussed above. This reaction is grounded in the entrenched budget culture, where programs compete with each other in an environment of changing rules, and where the issue of what is or is not working is an isolated question unrelated to how programs contribute to improving indicators of child and family well-being.

Trying to answer the wrong question, or The question of grain size: The question to be answered in results-based budgeting is: "What works to turn the indicator and cost-of-bad-results curves?" This is a very different question from "Does the educational system or welfare system or health care system work or not?" These are largely useless and unanswerable questions in a results framework. We want to know what works to turn the curves of children succeeding in school. If these curves are not turning, it is right to look at what is wrong with our strategies and to try to strengthen the system and its component parts. This will lead us to look at what strategies we think have a chance to turn the curves. But it will most often lead us to build on system components that are working, not to judge and abandon whole systems.

An axe to grind vs. agenda neutrality: This framework has no hidden agenda. It is about a simple process of being honest with ourselves, about trying things that we think can make a difference. It is not liberal or conservative. It is not about any particular ideology or program. We all bring with us convictions about what we think will work, and what can and should be done. But this process, like the current budget process, is a framework within which that debate can take place. The content of the debate is not dictated by the framework itself. Hopefully, those who prevail will help craft strategies that make the lives of children and families demonstrably better, and will have the wisdom and skill to see their strategies implemented and sustained. But nothing about this framework predetermines what strategies will emerge or whether they will or will not work.

Use of the phrase "what works" may require that we pay close attention to those whose reaction to the phrase is not grounded in an understanding of results-based systems. It would, of course be easy to overreact to this type of problem, and to replace each phrase that is subject to misinterpretation with some specially crafted piece of jargon. Common sense and clear language are essential ingredients for this work to succeed, but inadvertent or deliberate misinterpretation will always be with us.

G. The Departmental Trap:

Since results are broader than any one department, it is not clear what it means to try results-based budgeting within a single department. But that is how we generally organize work. And there is a temptation to designate some department or program to "try this first." This approach misses the point of results-based budgeting. The work must start with a result or a indicator that, by definition, spans departments or programs. In some cases, it may not be necessary to create a new structure. As noted above, many places already have interdepartmental groups working with charges that are, or could easily become, closely aligned with a result framework. These groups could use the results-based budgeting process to structure their work.

Still another misuse of departmental structure figures in the development of results and indicators. In one state, the Governor came away from a presentation of Oregon's benchmarks and soon after asked his cabinet heads to produce benchmarks for each agency. This approach

missed the point that goals and benchmarks in the Oregon framework by definition span department lines. One of the most daunting challenges in implementing results-based budgeting processes is countering this deeply entrenched tendency to think about government work in strictly traditional organizational terms.

H. Values and the Double Edge of the Cost of Bad Results:

The way we solve problems, the types of solutions which are acceptable, and the ways in which these solutions are developed and implemented all reflect the values of the community and the decision-makers. Nowhere is this more evident than in discussions of what works to turn the cost curves. The cost-of-bad-results analysis, if completed according to the design here, will include many, if not most, programs that can be viewed as society's charitable response to the needs of poor and disadvantaged families and children. There is a clear, simple, and uncharitable way to turn this curve by cutting these programs and allowing the needs to go unmet. Cutting these programs to save money is an expression of values.

The most important elements in any decision-making system are the values of the decision-makers. It should not be necessary to point out that these systems are as prone to misuse as any of the decision-making systems currently in place. For those whose values include a strong sense of social responsibility for children and families who are (temporarily or permanently) dependent on the good will of society, the point of the COBR picture is to find a "right" way to turn the curve, that is, by shifting spending to prevention and reducing the <u>need</u> for remedial social supports. Those who argue for direct cuts in social spending may achieve COBR reductions in the short term at the expense of increases in COBR expenditures in the long term. The challenge is how to cut these costs over the next 20 years. And the answer, over that period, must involve investments in prevention and support for children and families and that go beyond covering this year's deficit.

I. The Implications of a Results Approach for Cutting:

The challenge in government is more often how to cut, not increase, spending. What do results-based decision-making systems have to tell us about cutting expenditures? Results-based budgeting is about making the best decisions with available resources, whether those resources are growing or shrinking. Consider the following two approaches to cutting expenditures:

A traditional way of cutting: The traditional way in which decisions are made about cuts in human services can be depicted in a chart listing programs or functions in two columns and three rows. The columns distinguish between "mandated" and "non-mandated" programs– i.e., whether the program (or its functions) is required by law or regulation. Some clear examples of mandated programs include child abuse investigations, special education services, and child support enforcement services. As the federal government moves toward block grants and the states move toward managed care, the hard line between mandated and non-mandated services has begun to blur.

The second dimension is less a matter of law, and more a matter of judgment. To what extent is the service or function tied to the life, health, and safety of citizens? On this scale we can group programs or functions into three rough categories, reflecting low, medium, or high impact on life, health, and safety. A "high" on this list might include protective services investigations of child abuse, while a "low" might include administrative functions like contract management, or respite care services. In practice, we tend to cut programs and functions that are non-mandated, or perceived as having low impact on life, health, and safety. These tend to be the prevention programs that have no immediate safety implications, and the administrative infrastructure necessary to run programs and agencies. Conversely, we protect from cuts those services that are mandated, and rate high in impact on life, health, and safety. These tend to include the array of high-cost remediation programs.

This matrix goes a long way to explaining why and how we have disinvested in prevention over the last 10 years and why administrative capacity has been shaved back so severely in state and local government human services.

A results-based approach to cutting: Now consider an approach to cutting that uses a results framework as a starting point. Here we construct a two-column chart that distinguishes first between services that are matters of social infrastructure (or maintenance) like welfare payments (or highways), and second, programs that are devoted to "turning the curve" on results or indicators and the costs of bad results. For purposes of deciding on cuts, we judge each of these categories by separate criteria.

The first category (infrastructure/maintenance) we judge primarily on the basis of values. What level of bridge and highway disrepair are we willing to tolerate? What level of subsistence living are we as a society willing to adopt for our poorest children and their families (welfare benefit levels)? Obviously, there are facts that bear on these judgments, but the question at hand is principally a value-based judgment. We cut these programs on the basis of what we consider "essential" and what can be considered "non-essential."

With regard to the "turn-the-curve" part of the budget, the cutting criteria are easy to state but hard to answer. It is the same question of "what's working and what's not working" from the framework itself. Here decisions to cut focus first on what is not working, or what is not working as well as other parts of the strategy.[20] This process must be grounded in a willingness to honestly confront evidence of program ineffectiveness and to articulate a redirection agenda that reflects what it takes to improve results at scale.

J. The Relationship between Results and Performance Budgeting:

One of the most important lessons from our recent work on results-based budgeting is a fuller understanding of the difference between results and performance accountability. Performance accountability involves the use of performance measures to address whether

[20]It is, unfortunately, quite common to find examples of cuts in "turn-the-curve" programs that are clearly working. Lisbeth Schorr's book Within Our Reach documents a number of programs that have been defunded in spite of evidence of their effectiveness. It may be that a results-based decision-making framework, including a results-based approach to cutting, would make such actions less likely.

agencies and programs are working properly. This performance measurement process lies within a larger process that addresses whether our strategies, taken together, are working to turn the indicator and cost curves. This process-within-a-process perspective on the relationship between results and performance-based processes may be as important as the initial distinction between results, indicators, and performance measures.

Still another way to think about this difference is to consider the interrelated set of questions that frame result and performance accountability, such as:

1. Results: What do we want for our children, families, and communities?
2. Indicators: How do we know (how will we know) if we have achieved the results we want?
3. "What Works" Strategies: What (do we think) works to achieve the results we want?
4. Performance Measures: How do we know that the elements of our ("what works") strategy are performing as well as possible?

Results accountability addresses questions 1, 2, and 3. Performance accountability addresses questions 3 and 4. The two sets of questions are linked by the strategies developed to turn the indicator and cost curves. Each set of questions involves a feedback process that uses information on actual achievement to make strategy decisions (in the case of results and indicators) and management decisions (in the case of performance measures).

Results and performance accountability is often confused in practice, and this can be a debilitating confusion. When performance systems require agency directors to report on indicators, this creates the impression that agencies are accountable for indicators that are beyond their control. Agencies can be held accountable for their standing on agency and program performance measures. But improvements on results and indicators require that agencies act in partnership with other organizations and individuals inside and outside government, and this means that improvements cannot be the sole responsibility of a single agency.

Consider, for example, the rate of child abuse as an indicator of child health and safety, and the performance-measure percentage of child-abuse reports that are investigated within 24 hours. A state or local director of child welfare can be held accountable for the agency performance on response time. But impacting the incidence of child abuse is a "turn-the-curve" matter for the entire community. Unless the system in use establishes a clear distinction between results and performance accountability, these differences are easily lost, and the chances of making progress on either results or performance accountability is diminished.

Much, if not most, of the work under the banner of results accountability focuses primarily on agency and program performance measures, or fails to make this crucial distinction. State and local efforts that understand and can keep separate these important processes have a much better chance of making progress on both fronts.

V. CONCLUSION:

This paper begins and ends with results. If results for children are important, if improving them is important, and if we agree that what we have been doing is not working very well, then maybe it is time to try something different.

Results-based budgeting is a different, and we think promising, approach to making needed changes in our public budgeting systems. It is not an end in itself. And it is not a proven technology waiting to be replicated. Rather, it is a set of ideas ready to be tested by states and localities. The test in this case is whether this approach works to improve results. If results-based budgeting as presented here, or in some other form, cannot meet this single measure of success, then we should look to other options. Public budgeting systems, however, must soon begin to align the use of resources with our long-term social and financial interests in the well-being of children and families. It may turn out that failure to make this change is the most expensive of all our options.